



Genomic prediction of manganese efficiency in winter barley

Leplat, Florian; Jensen, Just; Madsen, Per

Published in:
The Plant Genome

DOI:
[10.3835/plantgenome2015.09.0085](https://doi.org/10.3835/plantgenome2015.09.0085)

Publication date:
2016

Document version
Publisher's PDF, also known as Version of record

Document license:
[CC BY-NC-ND](#)

Citation for published version (APA):
Leplat, F., Jensen, J., & Madsen, P. (2016). Genomic prediction of manganese efficiency in winter barley. *The Plant Genome*, 9(2). <https://doi.org/10.3835/plantgenome2015.09.0085>

Genomic Prediction of Manganese Efficiency in Winter Barley

Florian Leplat, Just Jensen, and Per Madsen*

Abstract

Manganese efficiency is a quantitative abiotic stress trait controlled by several genes each with a small effect. Manganese deficiency leads to yield reduction in winter barley (*Hordeum vulgare* L.). Breeding new cultivars for this trait remains difficult because of the lack of visual symptoms and the polygenic features of the trait. Hence, Mn efficiency is a potential suitable trait for a genomic selection (GS) approach. A collection of 248 winter barley varieties was screened for Mn efficiency using Chlorophyll *a* (Chl *a*) fluorescence in six environments prone to induce Mn deficiency. Two models for genomic prediction were implemented to predict future performance and breeding value of untested varieties. Predictions were obtained using multivariate mixed models: best linear unbiased predictor (BLUP) and genomic best linear unbiased predictor (G-BLUP). In the first model, predictions were based on the phenotypic evaluation, whereas both phenotypic and genomic marker data were included in the second model. Accuracy of predicting future phenotype, $r(\hat{g}, \bar{y})$, and accuracy of predicting true breeding values, $r(\hat{g}, g)$, were calculated and compared for both models using six cross-validation (CV) schemes; these were designed to mimic plant breeding programs. Overall, the CVs showed that prediction accuracies increased when using the G-BLUP model compared with the prediction accuracies using the BLUP model. Furthermore, the accuracies $[r(\hat{g}, g)]$ of predicting breeding values were more accurate than accuracy of predicting future phenotypes $[r(\hat{g}, \bar{y})]$. The study confirms that genomic data may enhance the prediction accuracy. Moreover it indicates that GS is a suitable breeding approach for quantitative abiotic stress traits.

Core Ideas

- The prediction accuracies of the G-BLUP model clearly outperformed the BLUP model
- Greenhouse experiments are useful information for Mn efficiency
- Prediction accuracies $r(\hat{g}, g)$ are more accurate than $r(\hat{g}, \bar{y})$
- Genomic prediction is a promising tool for plant breeding for nutrition traits

DESPITE that nutrient utilization in plants has been widely studied for many nutrients, the implementation of plant nutrient utilization traits in breeding programs is still not well developed even though nutrient use efficiency in plants is an important component of crop production economy (Baligar et al., 2007; Araus et al., 2008). Therefore, besides the main economical traits, nutrient use efficiency has become of major interest in recent breeding studies. Developing crops with higher nutrient use efficiency is necessary to counteract a range of worldwide plant nutritional disorders (Chapin et al., 2012) and the future rising costs and limited availability

F. Leplat, University of Copenhagen, Dep. of Plant and Environmental Sciences, Thorvaldsensvej 40, 1871 Frederiksberg C, Denmark; J. Jensen and P. Madsen, Aarhus Univ., Dep. of Molecular Biology and Genetics - Center for Quantitative Genetics and Genomics, Blichers Allé 20, 8830 Tjele, Denmark. Received 15 Sept. 2015. Accepted 22 Feb. 2016. *Corresponding author (per.madsen@mbg.au.dk).

Abbreviations: BLUP, best linear unbiased predictor; CV, cross-validation; Chl *a*, Chlorophyll *a*; EBV, estimated breeding value; F_v/F_m , ratio of variable fluorescence to maximum fluorescence; G-BLUP, genomic best linear unbiased predictor; GEBV, genomic estimated of breeding value; GP, genomic prediction; GS, genomic selection; GWAS, genome-wide association scan; MAS, marker-assisted selection; QTL, quantitative trait loci; SNP, single-nucleotide polymorphism.

Published in Plant Genome
Volume 9. doi: 10.3835/plantgenome2015.09.0085

© Crop Science Society of America
5585 Guilford Rd., Madison, WI 53711 USA
This is an open access article distributed under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

of fertilizers (Gilbert, 2009; Fixen and Johnston, 2012). Although, there is an increasing interest to select for nutrient use efficiency related traits to ensure a more sustainable use of limited resources especially for the macronutrients (N, P, and K). However, micronutrient deficiencies also have a direct impact on crop quality and yield (Hawkesford and Barraclough, 2011). Nutrient use efficiency is often highly polygenic, difficult to phenotype under field conditions, affected by exterior parameters (temperature, soil, light, etc.) and, therefore, is difficult to incorporate in practical breeding programs (Collins et al., 2008; Parry and Hawkesford, 2012).

Manganese deficiency remains an unsolved nutritional problem affecting crop production worldwide (Reuter et al., 1988; Welch et al., 1991; Hebborn et al., 2005; Yang et al., 2007), and breeding new genotypes with better Mn efficiency is challenging. Yield losses; suboptimal use of N, P, and water (Marschner and Marschner, 2012); and decrease in plant winter hardiness (Schmidt et al., 2013; Stoltz and Wallenhammar, 2014) are consequences of Mn deficiency. The deficiency is prevalent in areas with well-aerated, high-pH soils containing free carbonates and high organic matter content (Broadley et al., 2012). Phenotyping of Mn deficiency is complicated because there are no visible symptoms at the early growth development stages of the plants (Schmidt et al., 2013). However, besides being an activating cofactor of >35 enzymes in plants (Socha and Guerinot, 2014), the key role of Mn is its irreplaceable requirement in the oxygen-evolving complex of photosystem II. Manganese catalyzes the water-splitting reaction to produce oxygen and provides electrons for the photosynthetic electron transport chain (Nickelsen and Rengstl, 2013). This feature has been exploited to diagnose Mn efficiency by chlorophyll *a* (Chl *a*) fluorescence in winter barley (Hebborn et al., 2005; Schmidt et al., 2013). The quantification of photosynthetic efficiency by the quantum yield efficiency of photosystem II (ratio of variable fluorescence to maximum fluorescence [F_v/F_m]) (Govindjee, 2004) has been revealed to be a relevant parameter to measure the Mn efficiency of plants. This technique is therefore a relevant high-throughput phenotyping method allowing for phenotyping large-variety collections and genetic studies.

From a breeding perspective, very little is known on the Mn efficiency. Previous studies on barley genotypes cultivated in soil with low available Mn have shown genetic variation for the Mn efficiency and suggested a partial genetic control (Graham, 1988; Hebborn et al., 2005; Pedas et al., 2005). However, Mn pathways and homeostasis are not fully understood, and several physiological mechanisms have shown to be influenced by genetic variation: Mn uptake capacity (Pedas et al., 2005), stability and photochemical efficiency of the photosynthetic apparatus (Husted et al., 2009), and exudation of enzymes to the rhizosphere (George et al., 2014). Former marker-genetic studies of Mn efficiency measured as yield improvements in response to alleviating Mn deficiency proposed that Mn efficiency is controlled by a

single gene or locus (Mell; Pallotta et al., 2000). This locus has been validated in the biparental population Amagi Nijo \times WI2585 with two flanking restriction fragment length polymorphism markers on barley chromosome 4HS (Pallotta et al., 2003). Recently, Mn efficiency was screened using Chl *a* fluorescence, and genome-wide association scan (GWAS) has been performed testing the effect of single-nucleotide polymorphisms (SNPs). However, it has shown a relatively low broad-sense heritability of the trait (0.18–0.31) implying that quantitative trait loci (QTL) explained a limited portion of the genetic variance in the trait (F_v/F_m). The identification of many SNPs affecting F_v/F_m suggested that several candidate genes were involved, confirming the highly polygenic characteristics of the trait (data unpublished). All significant associations showed a small effect on F_v/F_m , indicating that the use of classic marker-assisted selection (MAS) in a breeding program might be difficult.

Indeed, MAS has been widely used in the last decades for selection of major genes affecting important traits in different crops (Collard and Mackill, 2008; Moose and Mumm, 2008). However, using marker information for complex traits controlled by many QTLs of small effects remains an issue (Heffner et al., 2009; Jannink et al., 2010). Therefore, GS has been proposed as a promising strategy for using genomic information in plant breeding programs and has been shown to outperform MAS (Heffner et al., 2011a,b). To run GS, statistical tools that allow calculating breeding values based on genomic information are needed. This process of predicting breeding values based on genomic information is termed genomic prediction (GP). Genomic selection was first introduced by Meuwissen et al. (2001) and it is now widely used in animal breeding (Hayes et al., 2009; Gao et al., 2012; Su et al., 2012b). Instead of using one marker at a time, GS use all available marker information simultaneously. Genomic prediction can be implemented as a unified SNP-BLUP model with random regression on SNP genotypes that estimates all marker effect at once (Heffner et al., 2009), also called ridge regression best linear unbiased prediction. The estimation of marker effects is then used to calculate the genomic estimated of breeding values (GEBVs) (Meuwissen et al., 2001). An alternative but mathematically equivalent method is G-BLUP proposed by VanRaden (2008), where the marker information is used to build a genomic relationship matrix (Habier et al., 2007; Goddard, 2009; Piepho, 2009). The benefit of GS lies in the possibility to predict genotypes for which no phenotypic information is available. Cross-validation techniques are often used to assess the accuracy of the prediction, where a validation population is predicted based on information from a training population. In a CV procedure, the phenotypic data from the validation population are left out, whereas the genotypic information is kept. The training population is comprised of the remaining data containing both genotypes and phenotypes (Heffner et al., 2009; Crossa et al., 2011). Prediction accuracy of the breeding value is often calculated as the Pearson

Table 1. Summary of experiment framework used for F_V/F_M screening (varieties across experiments are the same).

Abbreviation	Environment	Sowing	Phenotyping date	Population size	Number of replicates	Growth stage	Number of fluorescence measurements	Controls
GH13A	Greenhouse	13 November	13 December	248	2	3–4 leaves unfolded	1584	4 varieties \times 6 replicates
GH13B	Greenhouse	13 November	14 January	248	2	5 leaves to beginning of tillering	1584	4 varieties \times 3 replicates
KS11	Kristianstad	11 September	11 November	112	10	5 leaves to beginning of tillering	3840	15 varieties \times 3 replicates
LJ12	Lejre	12 September	13 May	233	2	End of tillering	1728	Not included
LJ13	Lejre	13 September	13 November	248	3	Beginning of tillering	2376	12 varieties \times 3 replicates
ST12	Saunte	12 September	13 May	233	2	End of tillering	1512	Not included

correlation coefficient between predicted breeding values and the observed phenotypic values for a subset of records where the phenotypes were left out when training the model (Zhong et al., 2009; Heffner et al., 2011b; Poland et al., 2012; Rutkoski et al., 2012).

In plants, recent studies presented encouraging prediction accuracy with the use of several GS models, where different experimental designs, number of markers, crops, and traits were used (Crossa et al., 2010; Heslot et al., 2012). In wheat (*Triticum aestivum* L.), genotyping-by-sequencing combined with GP has recently been studied for yield components. Lado et al. (2013) reported GEBV prediction accuracies up to 0.85 for 1000-kernel weight, a highly heritable trait, whereas Poland et al. (2012) reported prediction accuracies up to 0.38 for the same trait. These results reflect the differences in predicting GEBVs under different experimental designs. Moreover, prediction accuracies dropped dramatically when multi-environment data influenced by genotype \times environment interaction was considered (Lado et al., 2013).

The objectives of this study were (i) to develop BLUP and G-BLUP models for F_V/F_M value, (ii) to compare both models for their variance components and genotype \times environment interactions, and (iii) to evaluate the prediction accuracies of phenotypic-based predictions and marker-based predictions of F_V/F_M using six CV schemes. This study may provide some initial insights on the use of GS for an unexploited trait, Mn efficiency, and, more generally, serve as an example for abiotic stress, all of which are difficult to phenotype and are expected to become important in plant breeding in the near future.

Materials and Methods

Barley Collection

A collection of 248 winter barley varieties was screened for F_V/F_M . The collection is comprised of 112 genotypes from the ExBarDiv collection (European Research Area in Plant Genomics funded project, http://pgrc.ipk-gatersleben.de/barleynet/projects_exbardiv.php), with cultivars marketed over the last 60 yr. The collection was enriched with 18 double-haploid from Sejet Plant Breeding, 27 varieties from CRA-IECR (Consiglio per la Ricerca e sperimentazione in Agricoltura–Genomics Research Centre, Italy), and 92 commercialized varieties from European plant breeding companies. The collection

is a mix of 139 two-row and 109 six-row barley types. The study of the population stratification, by using *structure* software (Pritchard et al., 2000), revealed four subgroups of varieties. Two groups are comprised of two rows with Northern geographical origins, one group comprised six rows from mid-northern Europe, and finally, a last group is an admixture of two rows and six rows with southern geographical origins (Supplemental Fig. S1).

Genotypic Data

The whole population of winter-type barley varieties was genotyped with SNP markers by TraitGenetics GmbH using the Illumina iSelect 9k barley Infinium chip (Comadran et al., 2012). Allele calling was performed at TraitGenetics using their own software system. After removing markers with a minor allele frequency <0.01 and a call rate <0.75 , a final set of 5706 SNPs remained for analysis. Out of these, 4761 markers had a known chromosomal location on the physical barley map (International Barley Genome Sequencing Consortium, 2012).

Phenotypic Data

GH13A and GH13B were greenhouse experiments (Table 1) performed in 2013. To induce Mn deficiency, soil from Sweden (55.58°N, 14.05°E, Skåne area), known to induce Mn deficiency on winter crops (unpublished data), was collected and prepared to cultivate the winter barley plants. Manganese deficiency is often observed in loose and uncompact soil. Therefore, the soil was mixed with fine and coarse Leca (Mn neutral) and perlite to lighten the mix. In volume, proportions were 45, 25, 15, and 15%, respectively, of soil, perlite, fine Leca (2–6 mm), and coarse Leca (6–10 mm) and placed in 2-L pots. Temperatures were kept between 7 and 13°C and plants grew under natural light. Three plants of the same variety were sown in each pot and biological replicates were placed in a randomized complete block design (Table 1). Pots were watered by capillarity for 10 min every 3 wk.

KS11, LJ12, LJ13, and ST12 were field experiments (Table 1) performed in 2011, 2012, and 2013. The KS11 experiment was conducted in 2011 in Sweden (55.94°N, 14.19°E; Skåne area) on a field well characterized to induce Mn deficiency. LJ12 and LJ13 are two experiments completed in 2012 and 2013 (Lejre: 55.38°N, 11.57°E; Sealand), and ST12 (Saunte: 56.04°N, 12.30°E; Sealand) was sown in 2012. The fields in Saunte and Lejre are prevalent

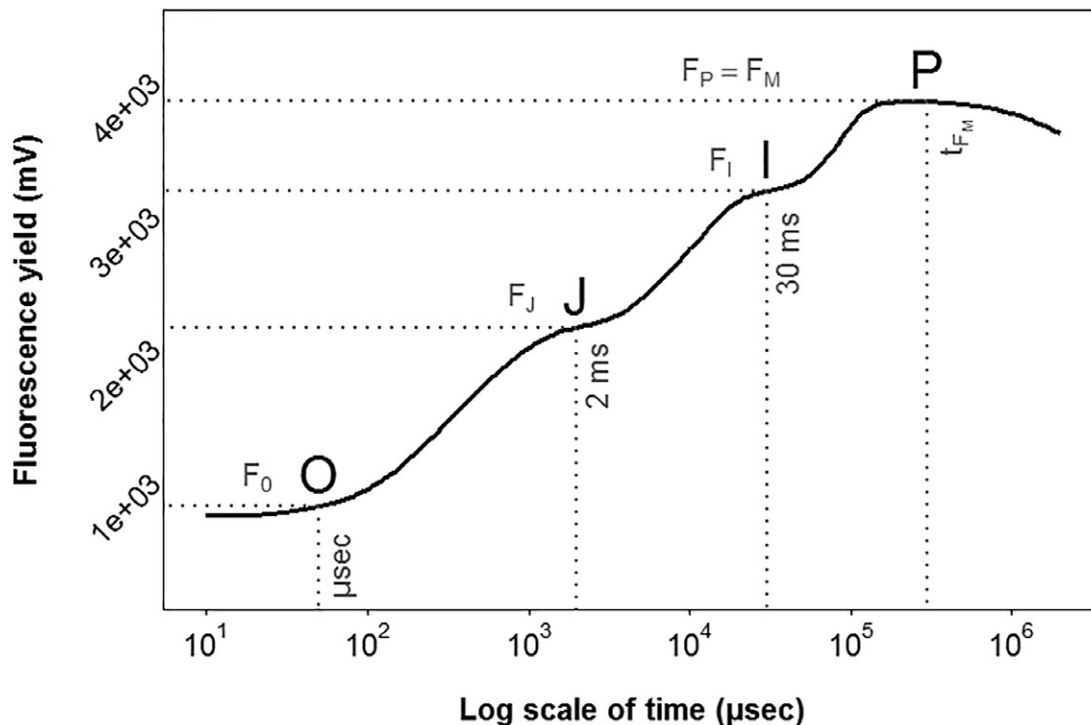


Fig. 1. A characteristic Chlorophyll *a* fluorescence of barley plant grown in optimal conditions. The horizontal axis represent the time on a logarithmic scale from 50 μ s to 1 s. The vertical axis characterizes the fluorescence yield in millivolts. The OJIP steps are highlighted. F_v/F_m is calculated following the formulae $(F_m - F_0)/F_m$.

Danish sandy soil type, where Mn deficiency had previously been observed. Field experiments were conducted in 1-m lines per variety and biological replicates were in a randomized block design (Table 1). Because of seasonal variation and operational limitations, screenings were done at different plant development stages (Table 1) and at temperatures $>5^{\circ}\text{C}$. In wheat and rye (*Secale cereale* L.), the acclimation at low temperatures is followed by a change in photosynthetic capacity (Rizza et al., 2001) and it would subsequently affect F_v/F_m measurements.

Phenotyping for Manganese Efficiency

Chlorophyll *a* fluorescence and its associated parameter F_v/F_m have been demonstrated to be a relevant and nondestructive method to diagnose Mn deficiency in winter barley (Hebborn et al., 2005; Husted et al., 2009). Schmidt et al. (2013) demonstrated that among Fe, Cu, S, and Mn deficiency, only Mn deficiency could induce a dramatic decrease of F_v/F_m . Moreover, in case of Mn resupply, F_v/F_m fully recovered until its maximum value. Therefore, F_v/F_m is a precise measure of the plant's ability to cope with Mn deficiency. The maximum quantum efficiency of photosystem II, the F_v/F_m value, was calculated as $(F_m - F_0)/F_m$ based on fluorescence transients (Fig. 1). This value is well known to characterize several abiotic stresses such as salinity, nutrition, freezing, and heating responses (Baker and Rosenqvist, 2004). A handheld plant efficiency analyzer device (Handy PEA, Hansatech Instruments Ltd.) was used to screen in vivo for Mn efficiency. Chlorophyll *a* fluorescence measurements were

conducted in the six environments: GH13A, GH13B, KS11, LJ12, ST12, and LJ13 (Table 1). The youngest fully emerged leaves were dark-adapted using Hansatech leaf clips for 25 min; Chl *a* fluorescence measurement were then measured for 2 s after illumination with a saturating light pulse of 3000 $\mu\text{mol photons m}^{-2} \text{s}^{-1}$ on the adaxial leaf surface. Three technical repeats were performed by measuring three different plants within each pot or line.

In each environment, control plants were sprayed with MnSO_4 applications (0.4% Mn solution containing two droplets of Tween 20 [Sigma-Aldrich] per liter) following the method described by Schmidt et al. (2013), and plants were measured 2 d later. A clear effect of MnSO_4 applications on F_v/F_m confirmed that the plants were Mn deficient at the time of measurement (Supplemental Fig. S2).

Statistical Models

The DMU package (Madsen and Jensen, 2013; Madsen et al., 2014) was used to perform the analyses using two models that were tested in six CV schemes. The ratio of variable fluorescence to maximum fluorescence (F_v/F_m) can be influenced by a number of factors such as temperature, root competition, or light exposition. These parameters could vary spatially over a field but also within greenhouses. Therefore, the models used included local environmental effects on individual pot and line and closely surrounding pots and lines. Each pot or line was considered to be influenced by the sum of nine local environmental spatial effects: one centered on its own

position and also affecting all its nearest neighbors and one effect centered on each of its eight closest neighbors, which in turn also affected all its nearest neighbors. The local spatial effect also captured the timescale frame used for the experiment because F_v/F_M measurements of pots and lines close to each other were recorded within a short time interval. Moreover, in the experimental design, one environment is considered as a combination of climate, plant growth stage, and controlled or field experiment.

Preliminary analysis showed that there was no genetic variation expressed in the two environments LJ12 and ST12 (Supplemental Table S1). They were therefore removed for the overall study. This could be explained by a mild and wet winter in 2012, whereas Mn deficiency often occurs with a relatively dry and cold climate. Consequently, the time of measurement was realized in spring, whereas Mn deficiency is often observed at early growth development stages. Therefore, because of the poor experimental phenotyping conditions, these two environments were left out in the analysis.

Finally, a four-variate mixed-model was used where each location is defined as subtraits. The model is described below:

$$y_{ijkrs} = \mu + X_i + Y_j + g_k + u_{ijk} + \sum_{f=1}^9 s_{ijkf} + e_{ijkrs}$$

where y_{ijkrs} is the phenotype of the observation placed in row i and column j of genotype k for the technical repeat r for one individual plant s within a pot or line; μ is the overall mean, X_i and Y_j are the fixed effects of coordinates, g_k is the random genotypic effect of variety k , u_{ijk} is the random effect of technical repeat (i.e., pot or line

effects), $\sum_{f=1}^9 s_{ijkf}$ are sum of local random spatial effects

estimated for squares centered on each pot or line and on each of the eight border pots or lines in the vicinity (therefore each pot or line is affected by a spatial effect centered on the pot or line and by spatial effects centered on each of the border pots or lines), and e_{ijkrs} is the residual. Coordinates of field plots were used to calculate the local effects; however, the plots constituting the border of the experimental field do not have eight surrounding plots. Therefore, to handle the local effect around those plots, virtual plots were added around them with missing phenotypes. The spatial effect of these virtual plots was estimated based on the observations in neighboring plots.

Variance components for both models were estimated using average-information restricted maximum likelihood algorithm as implemented in the DMU package (Madsen et al., 2014). The broad-sense heritability (H^2) of an individual measurement was computed according to Visscher et al. (2008). It reflects the contribution of the genetic variance ($\hat{\sigma}_g^2$) to the total phenotypic variance, with $\hat{\sigma}_g^2$, $\hat{\sigma}_R^2$, $\hat{\sigma}_L^2$, $\hat{\sigma}_e^2$ as the replicates, local spatial effects, and residual variance components, respectively. This definition returns the heritability of an individual measurement

on a random replicate of a random variety tested in a random spatial spot in the experiment.

$$H^2 = \frac{\hat{\sigma}_g^2}{\hat{\sigma}_g^2 + \hat{\sigma}_R^2 + 9\hat{\sigma}_L^2 + \hat{\sigma}_e^2}$$

Similarly, the heritability of the variety means (H_F^2) was also calculated using the following formula:

$$H_F^2 = \frac{\hat{\sigma}_g^2}{\hat{\sigma}_g^2 + \hat{\sigma}_R^2 / 10 + \hat{\sigma}_e^2 / 30}$$

The second definition of heritability assumes that every variety was tested using 10 biological replicates each measured using three technical replicates, as was the case for environment KS11. The same assumption was used for all experimental environments to ensure that the heritability estimates are comparable. Further, this definition assumes that inferences are made conditional on (corrected for) spatial effects.

Best Linear Unbiased Predictor Model

In the BLUP model, the genomic information is not included. Let \mathbf{g} be the vector of all genotypic effects, then $\text{var}(\mathbf{g}) = \mathbf{I} \otimes \mathbf{G}_0$, where \mathbf{G}_0 is the four-by-four matrix of genetic variances and covariances and \mathbf{I} is an identity matrix of order equal to the number of varieties. The subtraits were defined by location, therefore there are no residual covariances between subtraits, and the only covariances between locations are due to genetic effects. All random effects included were assumed to be multivariate normal distributed with zero mean.

Genomic Best Linear Unbiased Predictor Model

The principle of the G-BLUP model is similar to the multi-environment, marker-based model described by Burgueno et al. (2012), which models across-environment covariance matrices. Elements in the G-BLUP model are the same as in the former BLUP model, but the covariance matrix for genotypic effects were changed so that $\text{var}(\mathbf{g}) = \mathbf{G} \otimes \mathbf{G}_0$, where \mathbf{G} was the genomic relationship matrix, is calculated following the formula from (VanRaden, 2008):

$$\mathbf{G} = \frac{(\mathbf{M} - \mathbf{P})(\mathbf{M} - \mathbf{P})'}{\sum_{j=1}^m p_j(1 - p_j)}$$

where \mathbf{M} is a matrix of minor allele counts with m columns (m = total number of markers) and n rows (n = total number of genotyped varieties), and \mathbf{P} is a matrix containing twice the frequency of the minor allele (p_j). The \mathbf{G} matrix was setup using the Gmatrix program (Su and Madsen, 2011).

It was assumed, that \mathbf{G} includes the population structure information in such a way that the population structure contributes the predictive ability of the model.

Cross-Validation Schemes

Predicting the performance of new varieties that have not been yet been phenotyped is a fundamental issue in plant breeding. It allows deciding which genotype should be evaluated under real conditions or to obtain a prediction for varieties, which could not be phenotyped (like Mn efficiency). Cross-validation was used to assess the ability of the two models to predict future variety performances. A total of six across and within-environment CV schemes were designed. For each of these schemes, data were divided into a training population and validation population as reviewed by Heffner et al. (2009) and Crossa et al. (2011). The training population contained observations from individuals having phenotypic data (BLUP) or both phenotypic and genotypic data (G-BLUP). The validation population contained the remaining individuals of the dataset having no phenotypic data. The training population is then used to estimate model parameters, which are subsequently used to calculate the estimated breeding values (EBVs) and the GEBVs of the varieties in the validation population. Training population and validation population were randomly created to mimic real breeding programs where newly developed varieties or varieties in few environments had not been observed yet. In each CV scheme, 10 or 20% of the varieties were randomly left out to create the validation population, and the remaining varieties formed the training population. The GEBVs and EBVs were predicted using the two models. For each CV scheme, the procedure was repeated with 100 independent randomizations to create the training population. Prediction accuracies were calculated as an average over the 100 independent randomizations.

In CV1-10, 10% of the varieties that formed the validation population were from one environment only. Similarly, in CV1-20, 20% of the varieties were dropped and predicted in one environment. In these two schemes, the EBVs and GEBVs were predicted within environments by BLUP and G-BLUP models, respectively.

Similarly, in CV2-10 and CV2-20, respectively, 10 and 20% of the varieties were randomly left out in field experiments (KS11 and LJ13) to create the validation population and EBVs and GEBVs were predicted. In a breeding program, it corresponds to a case where phenotyping was not feasible in field but still possible in greenhouse controlled environments (GH13A, GH13B).

Finally, CV3-10 and CV3-20 are CV schemes where 10 and 20% of the varieties were randomly left out in all environments and breeding value predicted. It relates to a case, where a set of new varieties was not phenotyped in a breeding program. Therefore predictions for the BLUP model are not possible since no information is available for the varieties left out.

With the BLUP model, predictions for the different CVs come from correlated information, since the varieties were measured in other environments, whereas in the G-BLUP model, the prediction were based on the correlated traits (other environments) and the genomic

relationships matrix **G** among the varieties included in the data. Therefore, in the specific schemes CV3-10 and CV3-20, EBVs cannot be predicted using the BLUP model, whereas the genomic relationship matrix **G** allows predicting GEBVs using the G-BLUP.

For each CV scheme times replicate, EBVs and GEBVs were computed using variance components estimated with BLUP and G-BLUP models, respectively, on the full data as true parameters.

Prediction Accuracy

The prediction accuracy, $r(\hat{g}, \bar{y})$, is defined as the correlation between EBVs or GEBVs and the observed value corrected for the fixed effects of the validation population and averaged over the 100 replicates. The \hat{g} value is the predicted value (EBV or GEBV) of the validation population, and \bar{y} is the corresponding observed value corrected for the fixed effects. To correct for the fixed effect, the full model applied on all environments was run with the full dataset to estimate the fixed effects. Thus, the average of the correlations with 95% confidence interval was calculated over the 100 repeats of each CV schemes using the statistical environment R (R Core Development Team, 2014). The prediction accuracy, $r(\hat{g}, \bar{y})$, depends on the experimental design (number of records to compute \bar{y} for each variate). Therefore, accuracy of predicting breeding values was also computed as described by Su et al. (2012a):

$$r(\hat{g}, g) = \frac{r(\hat{g}, \bar{y}_c)}{r(\bar{y}_c, g)}$$

where $r(\bar{y}_c, g)$ is the square root of heritability of variety means but using the actual number of biological and technical replicates from each environment.

Results

Phenotypic Data

The phenotypic data was analyzed with both a multivariate BLUP and G-BLUP models, which take genotype \times environment interactions into account by modeling each environment as a separate subtrait. The variance components for all random effects were estimated in both models, and the contribution of the variance components to the total phenotypic variance was calculated (Fig. 2). The degrees of freedom (model complexity) of the models were the same, but a difference of 37 in the $-2 \log$ -likelihood value in favor of the G-BLUP model shows that the G-BLUP model had a better fit to the data. This was confirmed by the contribution of the variance components to the phenotypic variance. Results showed that the genetic variance component was slightly increased by the use of the genomic information in the G-BLUP model. This was the case for all four environments: GH13A, GH13B, KS11, and LJ13 (Fig. 2). Higher differences were observed for GH13A and KS11, where the heritability (H^2) of an individual measurement increased respectively from 0.29

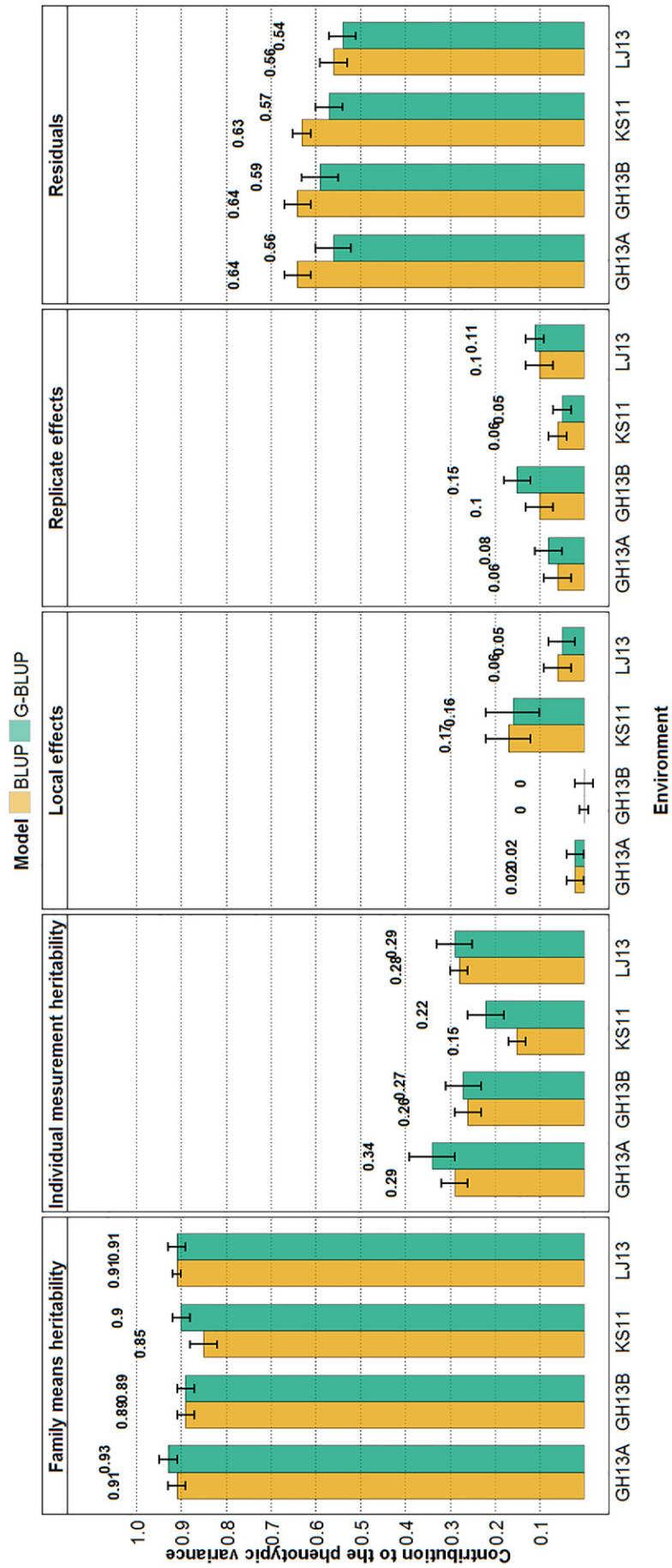


Fig. 2. Contribution of the variance components to the phenotypic variance. Heritability of individual measurement (H^2_i), heritability of variety means (H^2_v), local effects, replicates effects and residuals are estimated in best linear unbiased prediction (BLUP) and genomic BLUP models. The error bars show the standard errors, and numbers on top of the bars give the proportion of total phenotypic variance.

Table 2. Estimated phenotypic (σ_p^2) and genetic variance (σ_g^2); the contribution of random effect components and the genetic correlations between environments estimated by best linear unbiased prediction (BLUP) and genomic BLUP (G-BLUP) models for F_v/F_M trait. Standard errors are shown in brackets. In the table, σ_p^2 refers to the phenotypic variance used in the computation of H^2 .

Model	Environment	Variance component		Genetic correlation			
		σ_p^2	σ_g^2	GH13A	GH13B	KS11	LJ13
BLUP	GH13A	6.44×10^{-3} (2.95×10^{-4})	1.85×10^{-3} (2.70×10^{-4})	1			
	GH13B	1.33×10^{-2} (6.02×10^{-4})	3.47×10^{-3} (5.50×10^{-4})	0.767 (0.075)	1		
	KS11	6.62×10^{-3} (2.50×10^{-4})	9.64×10^{-4} (1.80×10^{-4})	-0.027 (0.144)	-0.101 (0.147)	1	
	LJ13	1.07×10^{-2} (4.51×10^{-4})	2.98×10^{-3} (3.96×10^{-4})	0.304 (0.090)	0.430 (0.089)	0.158 (0.138)	1
G-BLUP	GH13A	7.37×10^{-3} (4.59×10^{-4})	2.49×10^{-3} (4.66×10^{-4})	1			
	GH13B	1.45×10^{-2} (8.15×10^{-4})	3.88×10^{-3} (7.93×10^{-4})	0.892 (0.059)	1		
	KS11	7.26×10^{-3} (3.54×10^{-4})	1.58×10^{-3} (3.13×10^{-4})	-0.113 (0.156)	-0.189 (0.160)	1	
	LJ13	1.11×10^{-2} (5.77×10^{-4})	3.21×10^{-3} (5.59×10^{-4})	0.398 (0.111)	0.556 (0.102)	0.114 (0.153)	1

to 0.34 and from 0.15 to 0.22 (Fig. 2). Similar trends were observed for the heritability of the variety means (H_F^2) with noticeable differences for GH13A and KS11, respectively, from 0.91 to 0.93 and from 0.85 to 0.9. The heritability of variety means showed higher values, since it reflects the accuracy of using the mean of biological and technical replicates the F_v/F_M performance of a variety compared with the heritability of an individual measurement, which described the repeatability of a single F_v/F_M measure. The heritability of variety means indicated that the F_v/F_M screening was an accurate measurement with values higher than 0.85 in all four environments. Regarding the other parameters of the model, local effects were higher in field (0.05–0.16 for G-BLUP) than the greenhouse environment (0–0.02). This was expected, as greenhouses environment are better controlled than field environments. Finally, it was also necessary to consider the replicate effects of the F_v/F_M measurements, as they significantly contributed to the phenotypic variance (0.05–0.15). It can be explained by the Chl *a* fluorometer precision and also by the individual plants that were measured (three plants in one pot or line considered as one experimental unit).

Genetic correlations between environments (Table 2) were relatively low, ranging from -0.03 to 0.77 under the BLUP model and from -0.11 to 0.89 under the G-BLUP model, suggesting strong genotype \times environment interactions. Indeed, GH13A and GH13B, the controlled greenhouse environments, show the higher genetic correlation: 0.89 with G-BLUP vs. 0.77 with BLUP.

Prediction Accuracies According to Cross-Validation Schemes

The CV schemes have been designed to correspond to potential breeding situations. The CV1 scheme simulates a case where phenotyping data was missing in only one environment in a breeding program, but the phenotypes exist in the remaining environments. In CV1-10 and CV1-20, 10 and 20%, respectively, of the varieties were left out in only one environment at a time and constituted the validation population. It simulates the situation in a breeding program, where breeders have genotypes in

selection with missing information. It could be because of experiment size restriction, poor germination rate, technical limitations, occurrence of the symptoms, or for other phenotyping reasons. In that situation, parameters estimates from the BLUP and G-BLUP models are used to calculate EBVs or GEBVs based on the training population. It consisted of within-environment data and on correlated data across environments, respectively, without or with genomic relationship matrix. Overall, the prediction accuracy $r(\hat{g}, \bar{y})$ (Fig. 3) presented lower accuracy than the prediction accuracy $r(\hat{g}, g)$ (Fig. 4). The $r(\hat{g}, g)$ is the accuracy of predicting the true breeding value, which is more accurate than predicting future phenotypes. It is particularly the case when there were few replicates. The differences between $r(\hat{g}, \bar{y})$ and $r(\hat{g}, g)$ are higher for the predictions for GH13A and GH13B, which have less replicates (Table 1). For instance, in the G-BLUP model, in GH13A, $r(\hat{g}, \bar{y})$ for CV1-20 increased from 0.6 (Fig. 3) to 0.71 for $r(\hat{g}, g)$ (Fig. 4).

Both prediction accuracies $r(\hat{g}, \bar{y})$ and $r(\hat{g}, g)$ were always higher for the G-BLUP than for BLUP model in CV1-10 and CV1-20 for all environments (Fig. 3, 4). In controlled greenhouse for CV1-10, the accuracies were higher than in field experiment. For the G-BLUP model, $r(\hat{g}, \bar{y})$ (and $r(\hat{g}, g)$) in GH13A and GH13B were respectively 0.61 (0.72) and 0.56 (0.72), whereas prediction accuracies dropped to 0.3 (0.32) and 0.25 (0.29) in KS11 and LJ13 (Fig. 3, 4 [in parentheses]). The same pattern was observed for the prediction accuracies for the BLUP model; however, the BLUP model was completely ineffective in predicting LJ13. In that case, a large difference between G-BLUP prediction accuracies (0.25 [0.29]) and BLUP prediction accuracies (0.02 [0.02]) were revealed in CV1-10 (Fig. 3, 4). Overall, these results showed that for G-BLUP and BLUP models, CV1-10 and CV1-20, the greenhouse experiments were very consistent, with higher prediction accuracies than field experiments. Comparing both CV schemes, with a bigger validation population and consequently smaller training population, predictions were expected to be less accurate. Indeed, there was a decrease in prediction accuracies for CV1-20 in field experiment (Fig. 3, 4). However, prediction accuracies

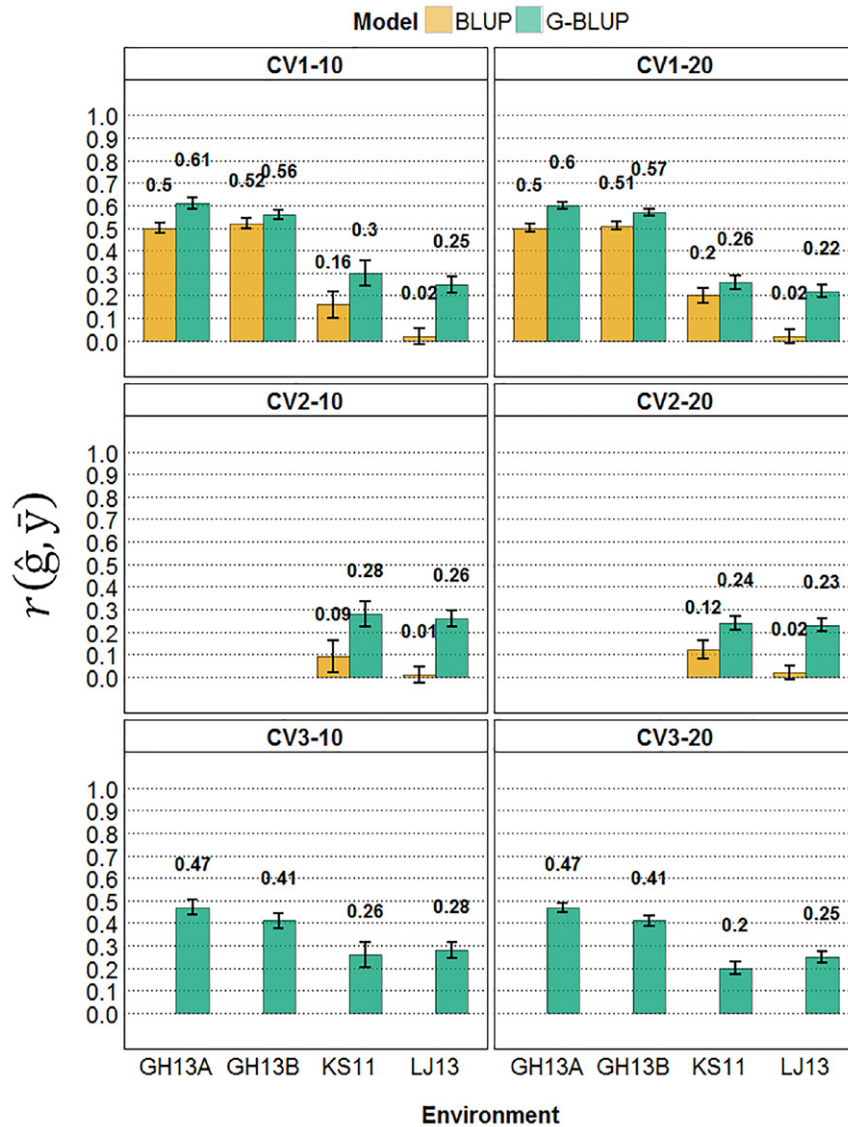


Fig. 3. Prediction accuracies of predicting future phenotypes, $r(\hat{g}, \bar{y})$, presented for best linear unbiased prediction (BLUP) and genomic BLUP models for the six cross-validation schemes analyzed. The numbers on top of the bars give the mean of prediction accuracy over the 100 cross-validation repeats. The error bar represents the 95% confidence calculated from a t -distribution.

were relatively stable in greenhouse experiments. Higher CV1-20 prediction accuracies for the BLUP model, in LJ13, vs. CV1-10 might be due to the randomized sampling of the validation population. Regarding CV2-10 and CV2-20, the breeding case corresponds to the situation where phenotyping is not achievable for some environments. Here, 10 and 20%, respectively, of the varieties were left out in both KS11 and LJ13 and constituted the validation population, simulating the situation where genotypes in selection have no field experiment information. Greenhouse experiments can most likely always be performed, unlike field experiments (KS11 or LJ13). Therefore, breeders could expect to predict the field performances based on correlated information between environments. Comparing the prediction accuracies $r(\hat{g}, \bar{y})$ [and $r(\hat{g}, g)$] in these CV schemes, G-BLUP, again, performed better than the BLUP model (Fig. 3, 4). Results led to similar statements from CV1-10 and CV1-20: BLUP model performed very poorly

in predicting EBVs in field experiments, with $r(\hat{g}, \bar{y})$ between 0.01 and 0.12, whereas G-BLUP produced $r(\hat{g}, \bar{y})$ values ranging between 0.23 and 0.28 and comparable to CV1-10 and CV1-20 (0.22–0.3) (Fig. 3). Prediction accuracies tend to decrease in field experiments when validation population size increased up to 20%.

Finally the last CV schemes, CV3-10 and CV3-20, characterize a situation where no information in any environment is known for a set of varieties. This is actually an ordinary situation in breeding programs. Indeed, depending on the breeding program scheme and the generation of selection, every year, breeders include new varieties for which no phenotypic information is known but for which genotypic information could be available. In that case, the interest of the G-BLUP model is evident because GEBVs could be calculated based on the genetic relationship matrix, whereas prediction of EBVs are not possible, since no information in any environment is known for

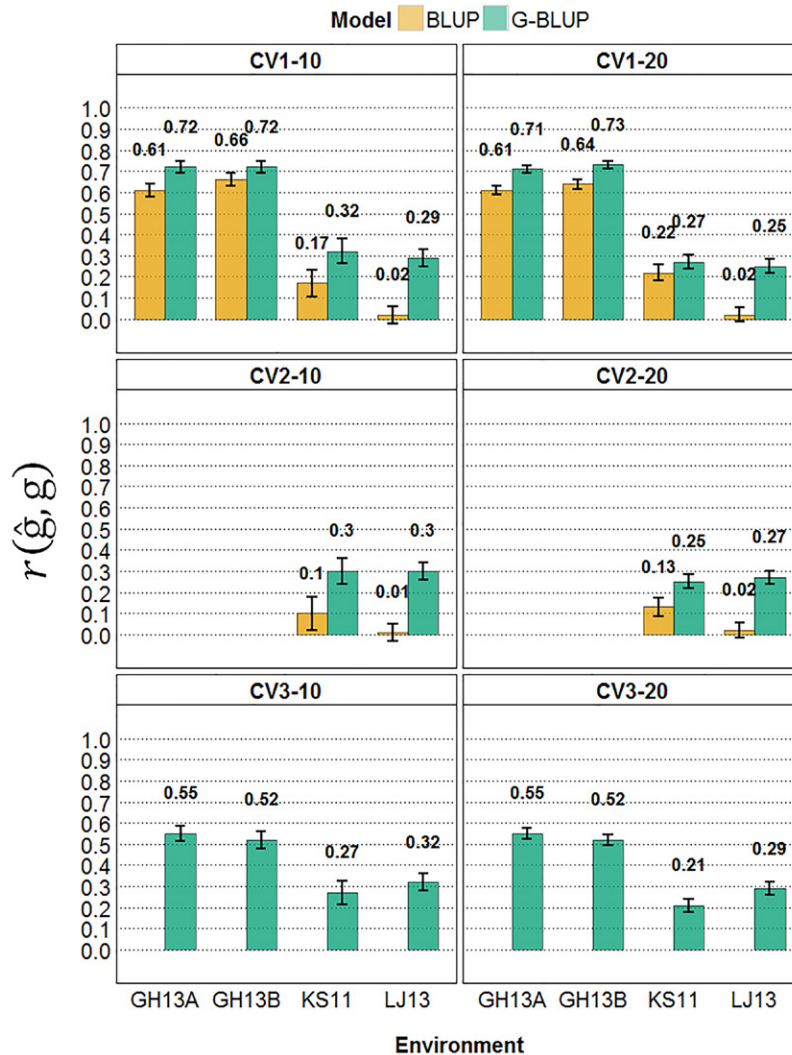


Fig. 4. Prediction accuracies of predicting true breeding values, $r(\hat{g}, g)$, presented for best linear unbiased prediction (BLUP) and genomic BLUP models for the six cross-validation schemes analyzed. The numbers on top of the bars give the mean of prediction accuracy over the 100 cross-validation repeats. The error bar represents the 95% confidence calculated from a *t*-distribution.

those varieties. Obtained results showed that greenhouse prediction accuracies were still higher than field experiments, and that the size of the validation population does not affect greenhouse prediction accuracies, whereas they decreased for field experiments (Fig. 3, 4). Moreover, similar to CV1-10 and CV1-20, greenhouse $r(\hat{g}, \bar{y})$ and $r(\hat{g}, g)$ values were stable between the two environments GH13A and GH13B and over the two CV schemes, with 0.47 (0.55) and 0.41 (0.52) in CV3-10 and CV3-20. Finally, despite a reduced training population in CV3-10 and CV3-20, r_M in these last two CV schemes were coherent compared with the four other schemes.

Discussion

Multivariate Mixed-Model Comparisons

Comparison of BLUP and G-BLUP model showed a clear advantage of using the G-BLUP model. The BLUP model is a standard method for estimating random effects of a mixed model, originally developed in animal breeding

(where pedigree information is often known), for estimation of breeding values. This model already revealed interesting characteristics for genotype predictions in plant breeding (Piepho et al., 2008). It can include the use of pedigree information, genetic correlation between traits, covariance between individuals, and a variance-covariance structure for modeling genotype \times environment interactions. However, the obtained results clearly confirmed that the BLUP model performance can be improved by the addition of the relationship matrix, **G**, based on marker information. The gain of G-BLUP model vs. the classical BLUP model has been suggested to be higher when traditional selection based only on phenotype is difficult (Goddard and Hayes, 2007). Case studies (Zhong et al., 2009; Lorenz et al., 2012) and a simulation study (Iwata and Jannink, 2011) in barley, based on the same markers and various GS models, clearly showed the applicability of GS in barley using SNP markers.

The present study confirmed the gain of the G-BLUP model designed for Mn efficiency. The G-BLUP model

seemed to be more suitable to capture a higher genetic variance and can be explained by modeling the covariance among varieties from the genomic relationship matrix. The heritabilities, the proportion of the genetic variances to the phenotypic variances, were slightly increased using the G-BLUP model in all environments analyzed, and results also showed a better fit of the G-BLUP model. According to both models, the experimental design showed strong genotype \times environment interactions with environments defined as the different experiments conducted (greenhouse vs. field). Overall, the G-BLUP model showed higher genetic correlations than the BLUP model. The high genetic correlation observed between the two greenhouse environments could be explained by the fact that both environments are very similar, with the growth stage as the main difference. With the G-BLUP, the relative genetic correlation of 0.56 between GH13B and LJ13, relating a controlled environment with a field environment LJ13, might be explained by the similar growth stage at the time of measurement. However, all genetic correlations with KS11 were null. It might be explained by the fact that the collection was not yet completed for KS11, and only 112 genotypes were screened for F_V/F_M in that environment. The overall low genetic correlation between field (KS11, LJ13) and the greenhouse experiments (GH13A, GH13B) can be expected, as the QTLs controlling the trait are most likely not stable across environments. Collins et al. (2008) introduced the concept of adaptive and constitutive QTLs. Constitutive QTLs are consistently expressed across environments, whereas adaptive QTLs would tend to be more environment specific. Therefore, for polygenic trait, the response in one environment would be determined by several adaptive QTLs of small effects not influencing other environments to the same degree.

Prediction Accuracies

Because of the large number of GP models and the various definitions of prediction accuracies, it is difficult to compare the results with other studies. Though $r(\hat{g}, \bar{y})$ and $r(\hat{g}, g)$ from the G-BLUP model showed clear accuracy gains over the BLUP model in all CVs, the accuracy of predicting the breeding value $r(\hat{g}, g)$ is more accurate than predicting a phenotype $r(\hat{g}, \bar{y})$. This is because the accuracy of predicting the phenotype is limited by the accuracy of the phenotype mean and this accuracy depends on the experimental design. Therefore, the accuracy of predicted breeding values show good potential to use genomic prediction instead of extensive phenotyping for complex traits such as Mn use efficiency.

In CV1-10 and CV1-20, the G-BLUP advantage is actually shown in field experiments, where the prediction accuracies were very poor; however, greenhouse accuracies were stable with a smaller gain for the G-BLUP model. In CV2-10 and CV2-20, the G-BLUP results were very similar to the two previous CV schemes, whereas the BLUP prediction accuracies dropped dramatically. Genomic selection seemed to be adequate to predict F_V/F_M

only in field environments based on training population containing the full greenhouse and part of the field phenotypic information. For a trait like Mn efficiency, where F_V/F_M screening is not possible every year under field conditions, controlled greenhouse experiments could, therefore, support field phenotyping and continue to increase and update the training population for future prediction. Finally, in the last two schemes CV3-10 and CV3-20, even though the overall G-BLUP prediction accuracies are slightly lower than in the other CV schemes, G-BLUP showed very stable characteristics with very similar prediction accuracies. The observed results demonstrated an important improvement of GEBV prediction using a G-BLUP model, leading to a conclusion comparable to the literature. Using molecular markers in G-BLUP, including modeling genotype \times environment interactions, gave higher prediction accuracy than not using molecular markers in a BLUP model (Burgueno et al., 2012). Simulation studies indicated that multi-environment prediction accuracy decreased as the proportion of shared QTL in the case of strong genotype \times environment interactions (Charmet et al., 2014). However, when using G-BLUP in multi-environment CV schemes, parameter estimates across environments allowed maintaining relatively constant prediction accuracy, whereas single-environment CV schemes were expected to present much higher prediction accuracies. In most of the CV schemes, only minor prediction accuracy decreases were observed when the validation population increased from 10 to 20%. It confirmed the robustness of the G-BLUP model toward changes in size of the training population. From the six CV schemes, it can be concluded that prediction accuracies from G-BLUP model clearly outperformed prediction accuracies of the BLUP model. Also, it showed the importance of having both field and greenhouse screening. Field experiments are closer to the real conditions but are less predictable. Prediction accuracies in greenhouse experiments were always higher than in the field experiments, but also the increased size of the validation population affected field experiments more than greenhouse experiments.

Perspectives for Selection for Manganese Efficiency

The progress of modern crop varieties is dependent on the improvement of highly polygenic traits, for which phenotypic selection is often an efficient method. Therefore, as a result of widespread soil fertility decline (Lal, 2009), combined with rarefaction and price increase of fertilizers, plant breeding has to develop new strategies to improve nutrient use efficiency traits. First, developments of high-throughput phenotyping tools are necessary especially for traits that are not easy to phenotype like nutrient deficiency. In the present case study of Mn efficiency, GS showed promising results for improving nutrient use efficiency compared with phenotypic selection only. As the phenotyping tools are available and the genotyping data becomes cheaper every year, the implementation of GS

for Mn efficiency will be possible. Nevertheless, breeders have to develop adequate strategies to build a powerful database including field experiments and maintenance of year-to-year greenhouse screening procedures.

Although, numerous applications of GS are expected in plant breeding, the next step would be to optimize the use of GS in practical breeding programs. It will involve standardized and affordable genotyping technologies and relatively simple breeding schemes designed for GS. Moreover, the characterization of the trait and experimental environments are required to handle genotype \times environment interactions. For example, for Mn efficiency, the use of a controlled environment helped to increase prediction accuracies. Finally, the size of the training and validation populations needs to be designed to ensure sufficient prediction accuracy.

Conclusions

The genetic background of Mn efficiency is highly polygenic and the difficulties to phenotype the trait under field conditions are limiting the development of new cultivars with high Mn efficiency. With developing marker technologies, the use of genomic data through GS can supplement classical phenotypic selection. The results obtained using BLUP and G-BLUP models showed that the G-BLUP model could be a relevant method to assist the breeders to select cultivars for Mn efficiency and similar traits. Indeed, use of the genomic relationship matrix substantially increased the prediction accuracies in all the six CV schemes used. The accuracy of predicting future performance using GEBV was stable across the different CV schemes and training and validation populations. In this case, with considerable genotype \times environment interactions, the G-BLUP model resulted in higher accuracy of predicting future phenotypes. A controlled environment, such as a greenhouse, could also provide useful data, allowing the breeder to update and increase the training population for future prediction. In a broader perspective, GS based on genomic relationships seem to be a suitable approach for selection for Mn efficiency. Further investigation of this approach combined with new high-throughput phenotyping method of other plant nutrition traits is warranted.

Acknowledgments

The authors acknowledge The Danish Strategic Research Council (DSF), which funded Nutriefficient project (Grant no: 10-093498) and GenSAP, Centre for Genomic Selection in Animals and Plants (Grant no: 0603-00519B12-132452). We thank Sejet Plant Breeding and Rasmus Lund Hjortshøj for their participation and for providing barley lines from their research program. The collection was provided by the participants of the ExBarDiv collection. We thank Laura Rossini and Alessandro Tondelli from University of Milan and Genomic Research Centre (Fiorenzuola), respectively, for the provision of seed and genomic data.

References

- Araus, J.L., G.A. Slafer, C. Royo and M.D. Serret. 2008. Breeding for yield potential and stress adaptation in cereals. *Crit. Rev. Plant Sci.* 27:377–412. doi:10.1080/07352680802467736
- Baker, N.R., and E. Rosenqvist. 2004. Applications of chlorophyll fluorescence can improve crop production strategies: An examination of future possibilities. *J. Exp. Bot.* 55:1607–1621. doi:10.1093/jxb/erh196
- Baligar, V.C., N.K. Fageria, and Z.L. He. 2007. Nutrient use efficiency in plants. *Commun. Soil Sci. Plant Anal.* 32:921–950. doi:10.1081/CSS-100104098
- Broadley, M., P. Brown, I. Cakmak, Z. Rengel, and F. Zhao. 2012. Function of nutrients: Micronutrients. In: P. Marschner, editor, *Marschner's mineral nutrition of higher plants*. 3rd ed. Academic Press, San Diego. p. 191–248.
- Burgueno, J., G. de los Campos, K. Weigel, and J. Crossa. 2012. Genomic prediction of breeding values when modeling genotype \times environment interaction using pedigree and dense molecular markers. *Crop Sci.* 52:707–719. doi:10.2135/cropsci2011.06.0299
- Chapin, F.S., III, P. Matson, and P. Vitousek. 2012. Plant nutrient use. In: *Principles of terrestrial ecosystem ecology*. Springer-Verlag, New York. p. 229–258.
- Charmet, G., E. Storlie, F.X. Oury, V. Laurent, D. Beghin, L. Chevarin, et al. 2014. Genome-wide prediction of three important traits in bread wheat. *Mol. Breed.* 34:1843–1852. doi:10.1007/s11032-014-0143-y
- Collard, B.C.Y., and D.J. Mackill. 2008. Marker-assisted selection: An approach for precision plant breeding in the twenty-first century. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 363:557–572. doi:10.1098/rstb.2007.2170
- Collins, N.C., F. Tardieu, and R. Tuberosa. 2008. Quantitative trait loci and crop performance under abiotic stress: Where do we stand? *Plant Physiol.* 147:469–486. doi:10.1104/pp.108.118117
- Comadran, J., B. Kilian, J. Russell, L. Ramsay, N. Stein, M. Ganai, et al. 2012. Natural variation in a homolog of *Antirrhinum CENTRORADIALIS* contributed to spring growth habit and environmental adaptation in cultivated barley. *Nat. Genet.* 44:1388–1392. doi:10.1038/ng.2447
- Crossa, J., G. de los Campos, P. Perez, D. Gianola, J. Burgueno, J.L. Araus, et al. 2010. Prediction of genetic values of quantitative traits in plant breeding using pedigree and molecular markers. *Genetics* 186:713–724. doi:10.1534/genetics.110.118521
- Crossa, J., P. Pérez, G. de los Campos, G. Mahuku, S. Dreisigacker, and C. Magorokosho. 2011. Genomic Selection and Prediction in Plant Breeding. *J. Crop Improv.* 25:239–261. doi:10.1080/15427528.2011.558767
- Fixen, P.E., and A.M. Johnston. 2012. World fertilizer nutrient reserves: A view to the future. *J. Sci. Food Agric.* 92:1001–1005. doi:10.1002/jsfa.4532
- Gao, H., O.F. Christensen, P. Madsen, U.S. Nielsen, Y. Zhang, M.S. Lund, et al. 2012. Comparison on genomic predictions using three GBLUP methods and two single-step blending methods in the Nordic Holstein population. *Genetics, selection, evolution. Genet., Sel., Evol.* 44:8. doi:10.1186/1297-9686-44-8
- George, T.S., A.S. French, L.K. Brown, A.J. Karley, P.J. White, L. Ramsay, et al. 2014. Genotypic variation in the ability of landraces and commercial cereal varieties to avoid manganese deficiency in soils with limited manganese availability: Is there a role for root-exuded phytases? *Physiol. Plant.* 151:243–256. doi:10.1111/ppl.12151
- Gilbert, N. 2009. Environment: The disappearing nutrient. *Nature* 461:716–718. doi:10.1038/461716a
- Goddard, M. 2009. Genomic selection: Prediction of accuracy and maximisation of long term response. *Genetica (The Hague)* 136:245–257. doi:10.1007/s10709-008-9308-0
- Goddard, M.E., and B.J. Hayes. 2007. Genomic selection. *J. Anim. Breed. Genet.* 124:323–330. doi:10.1111/j.1439-0388.2007.00702.x
- Govindjee, G. 2004. Chlorophyll a fluorescence: A bit of basics and history. In: Papageorgiou, G.C., Govindjee, editors, *Chlorophyll a fluorescence: A signature of photosynthesis*. Springer, Netherlands. p. 1–42.
- Graham, R. 1988. Genotypic differences in tolerance to manganese deficiency. In: R. Graham, R. Hannam, and N. Uren, editors, *Manganese in soils and plants*. Springer, Netherlands. p. 261–276.
- Habier, D., R.L. Fernando, and J.C. Dekkers. 2007. The impact of genetic relationship information on genome-assisted breeding values. *Genetics* 177:2389–2397. doi:10.1534/genetics.107.081190

- Hawkesford, M.J., and P. Barraclough. 2011. The molecular and physiological basis of nutrient use efficiency in crops. John Wiley & Sons, Chichester, West Sussex, UK.
- Hayes, B.J., P.M. Visscher, and M.E. Goddard. 2009. Increased accuracy of artificial selection by using the realized relationship matrix. *Genet. Res.* 91:47–60. doi:10.1017/S0016672308009981
- Hebborn, C.A., P. Pedas, J.K. Schjoerring, L. Knudsen, and S. Husted. 2005. Genotypic differences in manganese efficiency: Field experiments with winter barley (*Hordeum vulgare* L.). *Plant Soil* 272:233–244. doi:10.1007/s11104-004-5048-9
- Heffner, E.L., J.L. Jannink, H. Iwata, E. Souza, and M.E. Sorrells. 2011a. Genomic selection accuracy for grain quality traits in biparental wheat populations. *Crop Sci.* 51:2597–2606. doi:10.2135/cropsci2011.05.0253
- Heffner, E.L., J.L. Jannink, and M.E. Sorrells. 2011b. Genomic selection accuracy using multifamily prediction models in a wheat breeding program. *Plant Genome* 4:65–75. doi:10.3835/plantgenome.2010.12.0029
- Heffner, E.L., M.E. Sorrells, and J.L. Jannink. 2009. Genomic selection for crop improvement. *Crop Sci.* 49:1–12. doi:10.2135/cropsci2008.08.0512
- Heslot, N., H.P. Yang, M.E. Sorrells, and J.L. Jannink. 2012. Genomic selection in plant breeding: A comparison of models. *Crop Sci.* 52:146–160. doi:10.2135/cropsci2011.06.0297
- Husted, S., K.H. Laursen, C.A. Hebborn, S.B. Schmidt, P. Pedas, A. Haldrup, et al. 2009. Manganese deficiency leads to genotype-specific changes in fluorescence induction kinetics and state transitions. *Plant Physiol.* 150:825–833. doi:10.1104/pp.108.134601
- International Barley Genome Sequencing Consortium. 2012. A physical, genetic and functional sequence assembly of the barley genome. *Nature* 491:711–716. doi:10.1038/nature11543
- Iwata, H., and J.L. Jannink. 2011. Accuracy of genomic selection prediction in barley breeding programs: A simulation study based on the real single nucleotide polymorphism data of barley breeding lines. *Crop Sci.* 51:1915–1927. doi:10.2135/cropsci2010.12.0732
- Jannink, J.L., A.J. Lorenz and H. Iwata. 2010. Genomic selection in plant breeding: From theory to practice. *Briefings Funct. Genomics* 9:166–177. doi:10.1093/bfpg/eq001
- Lado, B., I. Matus, A. Rodriguez, L. Inostroza, J. Poland, F. Belzile, et al. 2013. Increased genomic prediction accuracy in wheat breeding through spatial adjustment of field trial data. *G3: Genes, Genomes, Genet.* 3:2105–2114. doi:10.1534/g3.113.007807
- Lal, R. 2009. Soils and food sufficiency: A review. In: E. Lichtfouse, M. Navarrete, P. Debaeke, S. Véronique, and C. Alberola, editors, Sustainable agriculture. Springer, Netherlands. p. 25–49.
- Lorenz, A.J., K.P. Smith, and J.L. Jannink. 2012. Potential and optimization of genomic selection for fusarium head blight resistance in six-row barley. *Crop Sci.* 52:1609–1621. doi:10.2135/cropsci2011.09.0503
- Madsen, P., and J. Jensen. 2013. DMU: A user's guide. A package for analysing multivariate mixed models. Version 6, release 5.2. University of Aarhus, Denmark.
- Madsen, P., J. Jensen, R. Labouriau, O.F. Christensen, and G. Sahana. 2014. DMU: A package for analyzing multivariate mixed models in quantitative genetics and genomics. 10th World Congress on Genetics Applied to Livestock Production, Vancouver, BC, Canada. 17–22 Aug. 2014. WCGALP, Auckland, NZ.
- Marschner, H. and P. Marschner. 2012. Marschner's mineral nutrition of higher plants. Academic Press, London.
- Meuwissen, T.H.E., B.J. Hayes, and M.E. Goddard. 2001. Prediction of total genetic value using genome-wide dense marker maps. *Genetics* 157:1819–1829.
- Moose, S.P., and R.H. Mumm. 2008. Molecular plant breeding as the foundation for 21st century crop improvement. *Plant Physiol.* 147:969–977. doi:10.1104/pp.108.118232
- Nickelsen, J., and B. Rengstl. 2013. Photosystem II assembly: From cyanobacteria to plants. *Annu. Rev. Plant Biol.* 64:609–635. doi:10.1146/annurev-arplant-050312-120124
- Pallotta, M.A., S. Asayama, J.M. Reinheimer, P.A. Davies, A.R. Barr, S.P. Jefferies, et al. 2003. Mapping and QTL analysis of the barley population Amagi Nijo × WI2585. *Aust. J. Agric. Res.* 54:1141–1144. doi:10.1071/AR02218
- Pallotta, M.A., R.D. Graham, P. Langridge, D.H.B. Sparrow, and S.J. Barker. 2000. RFLP mapping of manganese efficiency in barley. *Theor. Appl. Genet.* 101:1100–1108. doi:10.1007/s001220051585
- Parry, M.A., and M.J. Hawkesford. 2012. An integrated approach to crop genetic improvement. *J. Integr. Plant Biol.* 54:250–259. doi:10.1111/j.1744-7909.2012.01109.x
- Pedas, P., C.A. Hebborn, J.K. Schjoerring, P.E. Holm, and S. Husted. 2005. Differential capacity for high-affinity manganese uptake contributes to differences between barley genotypes in tolerance to low manganese availability. *Plant Physiol.* 139:1411–1420. doi:10.1104/pp.105.067561
- Piepho, H.P. 2009. Ridge regression and extensions for genomewide selection in maize. *Crop Sci.* 49:1165–1176. doi:10.2135/cropsci2008.10.0595
- Piepho, H.P., J. Möhring, A.E. Melchinger, and A. Büchse. 2008. BLUP for phenotypic selection in plant breeding and variety testing. *Euphytica* 161:209–228. doi:10.1007/s10681-007-9449-8
- Poland, J., J. Endelman, J. Dawson, J. Rutkoski, S.Y. Wu, Y. Manes, et al. 2012. Genomic selection in wheat breeding using genotyping-by-sequencing. *Plant Genome* 5:103–113. doi:10.3835/plantgenome2012.06.0006
- Pritchard, J.K., M. Stephens, and P. Donnelly. 2000. Inference of population structure using multilocus genotype data. *Genetics* 155:945–959.
- R Core Development Team. 2014. R: A language and environment for statistical computing. R Foundation for Stat. Comput., Vienna, Austria.
- Reuter, D., B. Cartwright, G. Judson, J. McFarlane, D. Maschmedt, and J. Robinson. 1988. Trace elements in South Australian agriculture. Technical Bulletin 139. Government of South Australia, Dep. of Agric., Adelaide.
- Rizza, F., D. Pagani, A.M. Stanca, and L. Cattivelli. 2001. Use of chlorophyll fluorescence to evaluate the cold acclimation and freezing tolerance of winter and spring oats. *Plant Breed.* 120:389–396. doi:10.1046/j.1439-0523.2001.00635.x
- Rutkoski, J., J. Benson, Y. Jia, G. Brown-Guedira, J.L. Jannink, and M. Sorrells. 2012. Evaluation of genomic prediction methods for fusarium head blight resistance in wheat. *Plant Genome* 5:51–61. doi:10.3835/plantgenome2012.02.0001
- Schmidt, S.B., P. Pedas, K.H. Laursen, J.K. Schjoerring, and S. Husted. 2013. Latent manganese deficiency in barley can be diagnosed and remediated on the basis of chlorophyll a fluorescence measurements. *Plant Soil* 372:417–429. doi:10.1007/s11104-013-1702-4
- Socha, A.L., and M.L. Guerinot. 2014. Mn-euvering manganese: The role of transporter gene family members in manganese uptake and mobilization in plants. *Front. Plant Sci.* 5:106. doi:10.3389/fpls.2014.00106
- Stoltz, E., and A.C. Wallenhammar. 2014. Manganese application increases winter hardiness in barley. *Field Crops Res.* 164:148–153. doi:10.1016/j.fcr.2014.05.008
- Su, G., O.F. Christensen, T. Ostensen, M. Henryon and M.S. Lund. 2012a. Estimating additive and non-additive genetic variances and predicting genetic merits using genome-wide dense single nucleotide polymorphism markers. *PLoS ONE* 7:e45293. doi:10.1371/journal.pone.0045293
- Su, G., and P. Madsen. 2011. User's guide for Gmatrix: A program for computing Genomic relationship matrix. University of Aarhus, Tjele, Denmark.
- Su, G., P. Madsen, U.S. Nielsen, E.A. Mantysaari, G.P. Aamand, O.F. Christensen, et al. 2012b. Genomic prediction for Nordic Red Cattle using one-step and selection index blending. *J. Dairy Sci.* 95:909–917. doi:10.3168/jds.2011-4804
- VanRaden, P.M. 2008. Efficient methods to compute genomic predictions. *J. Dairy Sci.* 91:4414–4423. doi:10.3168/jds.2007-0980
- Visscher, P.M., W.G. Hill, and N.R. Wray. 2008. Heritability in the genomics era: Concepts and misconceptions. *Nat. Rev. Genet.* 9:255–266. doi:10.1038/nrg2322
- Welch, R.M., W.H. Allaway, W.A. House, J. Kubota and R. Luxmoore. 1991. Geographic distribution of trace element problems. In: R.J. Luxmoore, editor, Micronutrients in agriculture. SSSA, Madison, WI. p. 31–57.
- Yang, X.E., W.R. Chen, and Y. Feng. 2007. Improving human micronutrient nutrition through biofortification in the soil-plant system: China as a case study. *Environ. Geochem. Health* 29:413–428. doi:10.1007/s10653-007-9086-0
- Zhong, S., J.C. Dekkers, R.L. Fernando, and J.L. Jannink. 2009. Factors affecting accuracy from genomic selection in populations derived from multiple inbred lines: A Barley case study. *Genetics* 182:355–364. doi:10.1534/genetics.108.098277